

Multi-level scalar structure in complex system analyses

Lipo Wang* and Guiwen Tan, Hui Cao

UM-SJTU Joint Institute, Shanghai JiaoTong Univ., Shanghai, China

(Dated: November 4, 2019)

The geometrical structure is among the most fundamental ingredients in understanding complex systems. Is there any systematic approach in defining structures quantitatively, rather than illustratively? If yes, what are the basic principles to follow? By introducing the concept of extremal points at different scale levels, a multi-level dissipation element approach has been developed to define structures at different scale levels, in accordance with the concept of structure hierarchy. Each dissipation element can be characterized by the length scale and the scalar variance inside. Using the two-dimensional fractal Brownian motion as a benchmark case, the conditional mean of the scalar difference with respect to the length scale shows clearly a power law and the scaling exponent is in agreement with the Hurst number. For the 3D turbulence velocity component, the 1/3 scaling law can be represented. These results indicate the important linkage between the turbulence physics and flow structure, if well posed and defined. In principle, the multi-level dissipation element idea is generally applicable in analyzing other multiscale complex systems as well.

Perhaps the most prominent feature of an organized system is its geometrical structure, either visualized or mathematically defined. Even for the systems with chaos [1] or deterministic chaos [2], the organized part may still co-exist with the dis-organized part. Geometry and shapes are definitely among the most fundamental ingredients in understanding complex systems. Considering turbulence for instance, ‘turbulent eddy’ is extensively used as an illustrative concept. Examples include the concept of energy cascade, stating that larger eddies pass energy to smaller eddies till the dissipative scale, and the β model for intermittency [3] with the hypothesis that the eddy volume decreases during the breaking process. Eddies of different sizes represent the coexisting multi-scales, both spatially and temporally. However, there is no quantitative definition of such kind of eddies or other spatial structures. Imaginary description is definitely unsatisfactory. Another implication of multi-scale is that the structure changes at different scale levels by ‘zooming’ the observation window. For instance, the spiral arm of the galaxy is the structure at the light-year level, while at the atomic scale level the bonding topology of the molecule appears. Self-similarity [4] is a special case that the appearance remains invariant or roughly the same on any scale. For the turbulence case, Corrsin [5] proposed the following questions to summarize the main challenges: “(1) What types (of geometry) are ‘naturally’ identifiable in turbulent flows? (2) What roles do they play or what properties do they have? and (3) What stochastic games can we invent which share some of the difficulties of the turbulent case, but are more treatable?”.

The existing structure analyses mainly suffer from two drawbacks: lack of quantification (for instance the vortex tube) or lack of non-local finite size characterization (for instance the critical points [6]). A systematic approach to define structures is by no means trivial. The concept of dissipation element (DE) put forward by Wang

& Peters [7] shed light on this problem. Starting from each spatial point in a turbulent scalar field, one will inevitably reach a local minimum and local maximum points of this scalar along the descending and ascending directions of the scalar gradient trajectory, respectively. The ensemble of spatial points from which the same pair of minimum and maximum points are reached define a spatial region, called a dissipation element. The unique and favorable features of the DE structure are summarized as follows. First, the structure can be parameterized to ensure quantitative analyses. In the existing work, the characteristic parameters of DE are l and $\Delta\phi = \phi_{max} - \phi_{min}$, the linear distance between two extremal points and the absolute value of the scalar difference, respectively. Second, DEs are space-filling, which enables the reconstruction of the entire flow field from the statistics of individual units. For instance, if the topological features of the decomposed units can be described by a parameter set $(p_1, p_2, \dots)$, in principle it is much easier and more accurate to represent a field property X by p_i within individual units than to construct X with respect to the entire field. Once the joint probability density function (PDF) of p_i , $P(p_1, p_2, \dots)$, has been modeled, the ensemble average of X , denoted by $\langle X \rangle = \langle X(p_1, \dots, p_n) \rangle$, is then determined by

$$\langle X \rangle = \int \dots \int X(p_1, \dots, p_n) P(p_1, \dots, p_n) dp_1 \dots dp_n. \quad (1)$$

By decomposing the entire field into sub-units, the system complexity may be reduced by understanding the unit statistics. The above relation is valid if and only if the units are space-filling. The widely used vortex tube structure based on the Q criterion, for instance, depends on the preset iso-values with arbitrariness; meanwhile, vortex tubes take only a small portion of entire field, especially at the higher Reynolds numbers.

METHODS

Considering an artificial field as shown in Fig. 1 (a), two different scale structures appear: the small scale ripples and the large scale wave. Both structures need to be recognized, if the structure definition is well posed. Because the scalar gradient trajectories stop at local extremal points, the DE approach can only capture the ripple-like units. In reality, data noises are inevitable to largely alter the spatial distribution of extremal points and DEs as well. DEs constructed in this way can be considered to image the turbulent eddy entities at the fine scale level, but the structures at higher levels are missing. An ideal structure representation should not be sensitive to data noises or small changes of appearance; otherwise, similar objects will be represented quite differently.

A satisfactory remedy to overcome this difficulty is the multi-level segment idea developed by Wang & Huang [9]. The framework is built upon a simple fact that extremal points are conditional valid, with respect to the preset observation window. The implication that a point x_0 is extremal at the r scale level is

$$\begin{cases} f(x_0) \geq f(x), \forall x \in |x - x_0| \leq r, (\text{maximum}) \\ f(x_0) \leq f(x), \forall x \in |x - x_0| \leq r, (\text{minimum}). \end{cases} \quad (2)$$

The distance measure $|\cdot|$ can be tailored according to the questions under consideration, e.g. the typically used Euclidean distance. In this sense, the local extremal (maximum or minimum) points are the extremal points at the scale level $r \rightarrow 0$, or in simplicity $r = 0$.

In one-dimensional space for illustration, Fig. 1 (b) shows all the extremal points (blue for minimum and red for maximum), which, however, are valid at different scale levels. Specifically, points marked in square have scale smaller than that of points in circle. For each given scale r , the set of corresponding extremal points can be extracted. A segment at the r level is defined as the part between two consecutive extremal points (at the r level). Typically, the segments become larger with the increase of r . Thus structures at different observation levels (i.e. different r) can be constructed. The multi-level segment approach [9] based on this principle proves to be effective to separate the mixed statistics at different scales. In the present work, the conditional validity defined in Eq. (2) will be used to extend the DE structure to multi-levels for a better understanding of shapes in complex data analyses. The detailed definition and algorithm are described as follows.

In Fig. 1 (c), starting from any spatial point P in a scalar field $f(\vec{x})$, the scalar gradient trajectory connects its local minimum point $B1$ and local maximum point $R1$. For a specified scale r , if $B1$ is still the minimum point at the r level, then the trajectory along the descending direction stops at $B1$. If $R1$ is not the maximum point at

the r level, it implies that there is another point $R2$ satisfying $f(R2) > f(R1)$ within the spherical domain centered at $R1$ with a radius r . Then the trajectory needs to jump to $R2$; if $R2$ is not yet a maximum point at the r level, then the trajectory jumps from $R2$ to $R3$. Such jump process continues till Rn , a maximum point at the r level. In other words, at the r level the gradient trajectory of point P now connects $B1$ and Rn . The gradient trajectory between $B1$ and $R1$ is the case for $r = 0$. In this sense, typically gradient trajectories at r need not to be continuous. A natural extension of DE at the r level is the set of all spatial points whose gradient trajectories share the same pair of minimum and maximum points at the r level. Let r scan from 0 to some large enough quantity. Then multiple level DEs are determined. Because of the confined influence range of noises, typically only the small structures will be strongly influenced by noises, while the large ones remain unacted. In other words, the noise influence can be easily removed by setting the scale level r above the noise range. Thus the multi-level DE concept functions even for noise contaminated data.

RESULTS

Benchmark tests

From the definition of the multi-level DE structure, it is ready to conclude that the decomposed units are space-filling at any r level, because each spatial point has its unique corresponding gradient trajectory. Reconsider the multi-scale structure as shown in Fig. 1 (a). The small ripple ones can be detected when r is small; while the large waves can also be extracted using large r . Similarly as the $r = 0$ level case, the multi-level DE can also be characterized by l and $\Delta\phi = \phi_{max} - \phi_{min}$, but the quantities at the r level. The joint probability density function (PDF) of $\Delta\phi$ and l , $p(\Delta\phi, l)$, is shown in Fig. 2 (a). The two sample clustering zones A and B justifies the scale separation, i.e. small scale elements contribute the samples in zone A , while large wave elements correspond to the samples in zone B . $p(\Delta\phi, l)$ links the statistics of the decomposed units and the entire flow, as indicated by Eq. (1). For case validation, we first analyze the multi-level DE structure of the two-dimensional fractional Brownian motion (fBm) field. Introduced by Kolmogorov [10] and extensively studied by Mandelbrot et al. [11], fBm has been considered as a classical scaling stochastic process in many fields [12–14]. A fast Fourier transform based on the Wood-Chan algorithm [15] is used to synthesis the data. The conventional structure function (SF) calculates that for any two spatial points separated with a scale x , the statistical mean of the field variable difference in between. Because of self-similarity, theoretically the scaling exponent $\zeta(q)$ of SF depends linearly on the moment order q , i.e.

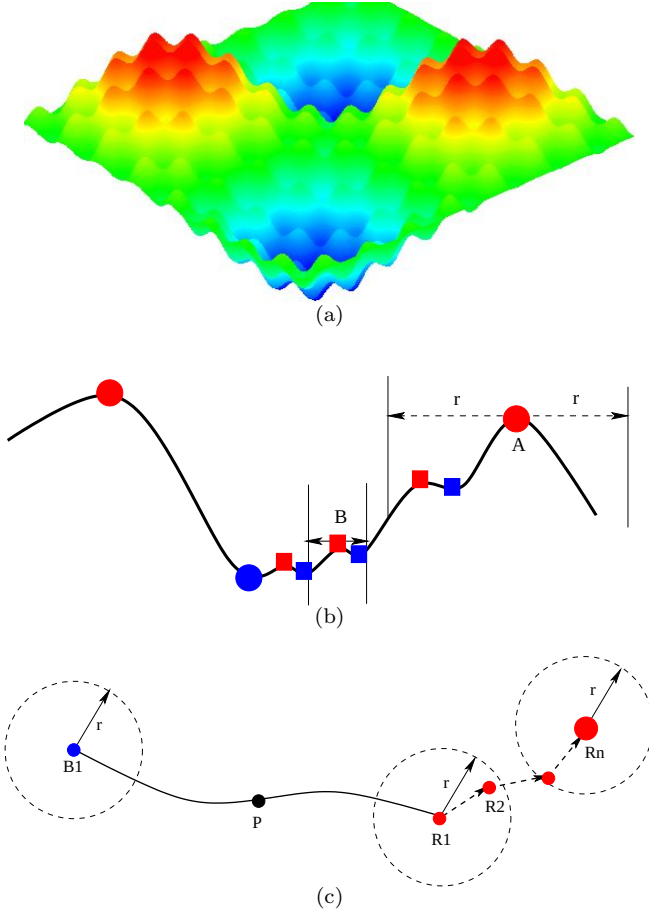


FIG. 1: (a) A scalar field (colored with the local value) with two different scale structures: the small scale ripple and the large scale wave. Local extrema are sensitive to data noises and small perturbations. Therefore DE analysis is ineffective to extract the large wave structure. (b) Extremal points are conditionally valid. If the window size is small enough, square and circle points are extremal (red for maximum and blue for minimum), while at some large window size, circle points are extremal but the square ones are not. Such property is closely related to the hierarchical structure of turbulence. (c) The algorithm designed to detect the ‘multi-level’ gradient trajectory. Starting from any spatial point P , the gradient trajectory connects one maximal point (small red dot) and one minimal point (small blue dot). For a specified scale r , if an ending point (e.g. point B_1) is still extremal with respect to a spherical domain with the radius of r , then the searching process stops; if not (e.g. the point R_1), then the gradient trajectory jumps to R_2 , the maximum within the r -sphere of the point R_1 . Such jumping process continues until R_n , which is the maximum within its r sphere.

$\zeta(q) = qH$, in which H is the Hurst number. In the context of DE, scale is determined by structure, but not an independent input. The conditional mean of $\Delta\phi$ on l , i.e. $\langle\Delta\phi|l\rangle = \int \Delta\phi P_{\Delta\phi}(\Delta\phi|l)d(\Delta\phi)$ can be interpreted as a newly defined first order SF with $q = 1$. Such definition introduces the influence of structure; meanwhile

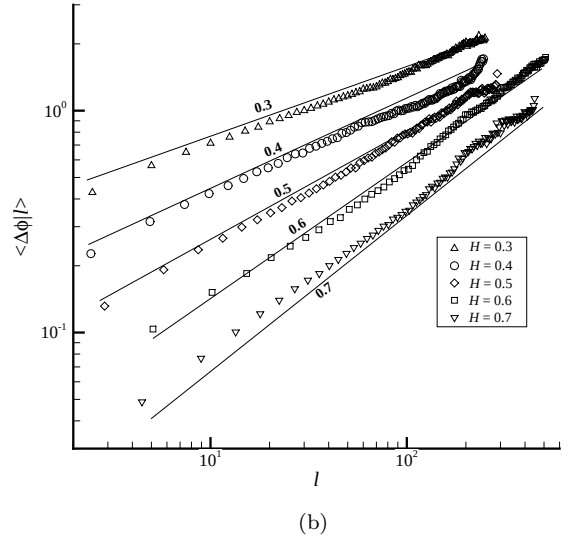
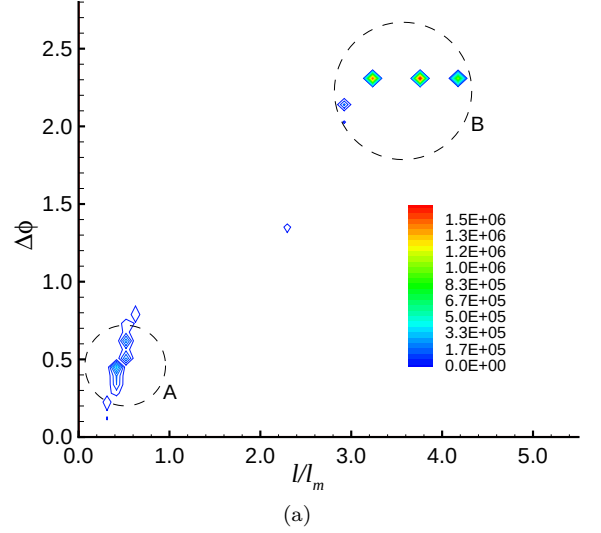


FIG. 2: (a) The joint PDF of $\Delta\phi$ and l , $p(\Delta\phi, l)$ corresponding to the multi-level DE statistics of the Fig. 1 (a) case. The two sample clustering zones A and B justifies the scale separation, i.e. small and large scale elements contribute the samples in zone A and B, respectively. (b) Scaling detected from the two dimensional fbM process with different Hurst numbers using the multi-level DE structure approach. The scaling exponent of $\langle\Delta\phi|l\rangle$ agrees satisfactorily with H .

it is more effective to separate the so-called scale-mixing of different correlation regions [7]. The results are shown in Fig. 2 (b). Clearly the scaling exponent of $\langle\Delta\phi|l\rangle$ from the multi-level DE structure agrees satisfactorily with theoretical prediction, i.e. the slope is equal to H . Because of self-similarity, this scaling is valid in the entire scale range.

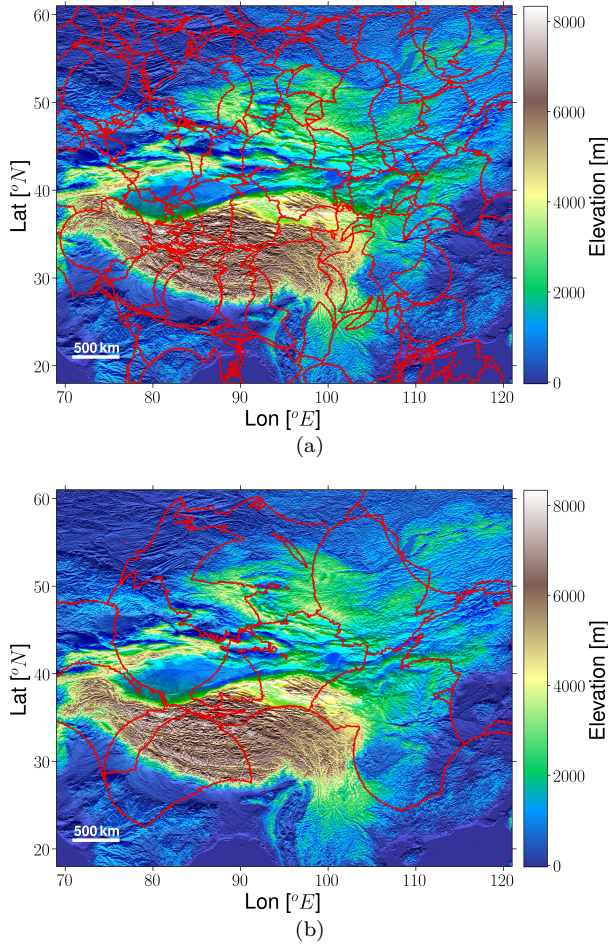


FIG. 3: Application of Multi-level DE analysis to the topographic data of the Tibetan plateau region (resolved with 2000×2000 grid points) at different observation window size r : (a) $r \sim 350\text{km}$ and (b) $r \sim 700\text{km}$. The red lines denote the DE boundaries. Clearly, when the window size doubles, the number of DEs decreases drastically and the size of DEs increases.

Data analyses: topographical structure

The multi-level DE approach is also implemented to analyze the topography data. Based on the model data of Earth's surface that integrates land topography and ocean bathymetry [16], the Tibetan plateau region resolved with 2000×2000 grid points is analyzed. DEs at two different window sizes are shown in Fig. 3 (a) ($r \sim 350\text{km}$) and (b) $r \sim 700\text{km}$. The red lines denote the element boundaries. Clearly, when the window size doubles, the number of DEs decreases drastically and the size of DEs increases. The irregularity of the earth surface leads to unsmoothed boundaries. According to the definition of multi-level DE, such partition is space-filling.

Data analyses: turbulence

Three-dimensional turbulence velocity field is also investigated. Direct numerical simulations are performed within a 2π cubic box for isotropic incompressible turbulence. The Reynolds number based on the Taylor scale is about 100. The dissipative scale can be sufficiently resolved using a 512^3 grid point mesh. First, examples of dissipation elements at different scale levels are shown in Fig. 4 (a). For the convenience of visualization, a two-dimensional slice from three-dimensional field is plotted. We choose two different r levels: $r = 0$ and 20 times grid space Δx . At the $r = 0$ level, element boundaries are shown in thin solid lines. At the $r = 20\Delta x$ level, the boundary of each DE is shown in thick yellow lines. Within each of these larger DE, there are number of $r = 0$ level elements, whose boundaries (in thin solid lines) are presented with same color. Typically the higher scale level structure looks more complex. For instance, extremal points can be located inside DE. As aforementioned, all the elements at the same r level are space-filling, which means the entire flow field can be decomposed differently when r changes. Differently from the Fig. 3, because of the molecular diffusivity, the scalar field in turbulence is fine-scale smooth and thus the boundaries are smooth as well.

Collectively, the joint PDF of $(\Delta\phi, l)$, $p(\Delta\phi, l)$, from the DE samples at different r are shown in Fig. 4 (b). It has been discussed [7, 8] that the evolution of extremal points are under the control of two counteracting mechanisms. On the one hand, turbulent random motions generate extremal points by disturbing the flow field; on the other hand, molecular diffusion annihilates closely clustered extremal points and small elements to smooth the field. Therefore the number density of extremal points remain unchanged in statistically stationary turbulence. This scenario remains valid at different r levels, making the joint PDF peaks at nonzero l and nonzero $\Delta\phi$. The conditional mean extracted from this joint PDF is shown in Fig 4 (c) in the log-log plot. The nice scaling in a broad range of l is consistent with Kolmogorov's scaling prediction, namely a $1/3$ power law. This result indicates that multi-level DE analysis is enlightening and meaningful in understanding turbulence .

In summary, geometrical analysis is important to understand turbulence and other complex systems, in the aspects of the field structure, kinematic and dynamic properties. For a given scalar field, the entire space can be decomposed into dissipation elements (DE), by tracing along scalar gradient trajectories till the local extremal points are reached. Such decomposition is non-arbitrarily defined and space-filling, which makes possible to understand the original entire field via the statistics of the decomposed units, as suggested by Eq. (1). The length scale of DEs is determined by the structure, but not a

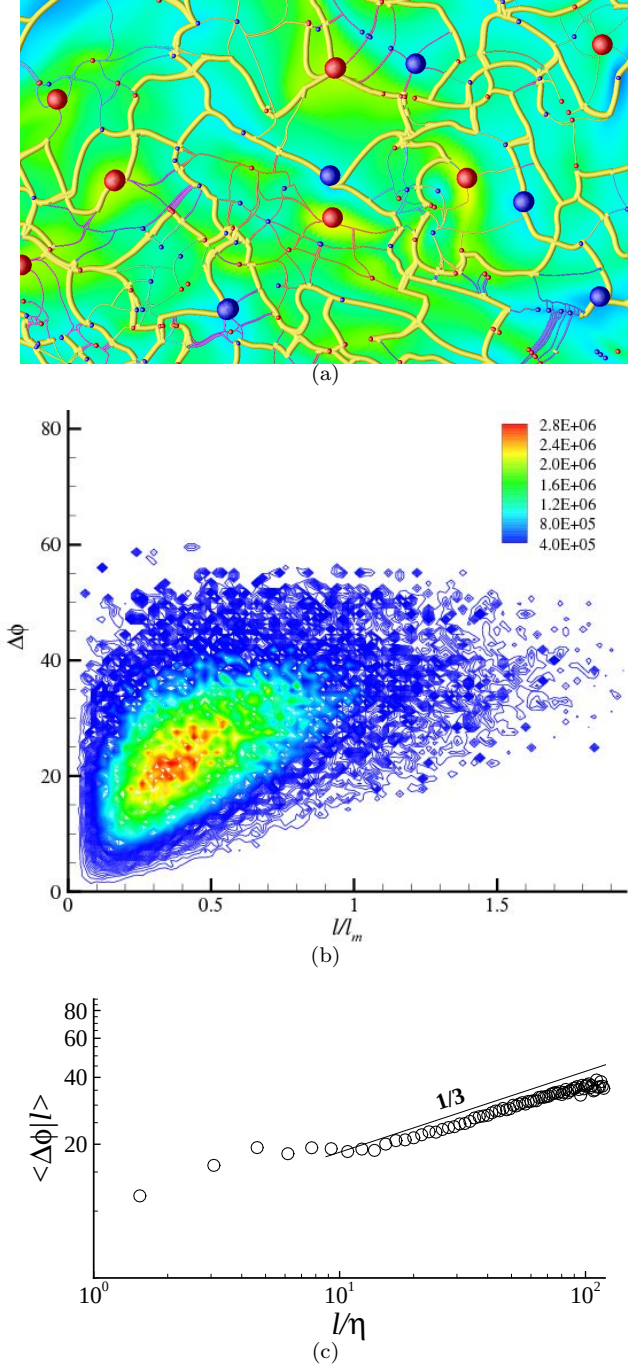


FIG. 4: Multi-level DE results from the three-dimensional isotropic turbulence DNS data: (a) Visualization of DEs at different scale levels (by processing a two-dimensional slice from three-dimensional isotropic turbulence DNS data). At the $r = 0$ level, the DE boundaries are shown in thin solid lines. The $r = 0$ level extremal points are marked as small blue and small red points. At the $r = 20\Delta x$ scale level, the DE boundaries are presented in thick yellow lines with the corresponding extremal points in large blue and large red dots. Each larger element (at $r = 20\Delta x$ level) encompasses some smaller ones (at $r = 0$ level) with a same boundary color; (b) the joint PDF of $\Delta\phi$ and l ; (c) the conditional mean of $\Delta\phi$ with respect to l , showing a $1/3$ scaling in agreement with the dimensional argument prediction.

independent parameter. In spite of the meaningfulness in data analysis, the DE structure is sensitive to noise contamination and fails to detect large structures effectively. A well posed remedy to overcome these deficiencies is that by introducing the concept of extremal point at the r scale level, the DE structure can be extended to multi-levels. For a specified scale r , the gradient trajectory of any spatial point need to jump starting from the 0 level extremal points till the r level extremal points are reached. Similarly a multi-level DE is the set of spatial points whose gradient trajectories share the same pair of extremal points at the r level. The characteristic parameters of DE are chosen as the distance l and the scalar difference $\Delta\phi$ between two extremal points. The conditional mean $\langle \Delta\phi | l \rangle$ can be interpreted as a newly defined structure function. For benchmark test, the fractional Brownian motion case is analyzed. It shows $\langle \Delta\phi | l \rangle \propto l^H$, a nice scaling law in agreement with the expectation. For the velocity component u field in 3D turbulence, in a broad scale range the conditional mean of Δu follows the dimensional argument scaling, i.e. $\langle \Delta u | l \rangle \propto l^{1/3}$. These results indicate that the multi-level DE structure does reveal the flow physics at different scale levels (or the observation window sizes). In principle, this non-arbitrarily defined and space-filling structure can be implemented to analyze other complex systems as well.

The authors acknowledge Dr. Y. Huang (State Key Laboratory of Marine Environmental Science & College of Ocean and Earth Sciences, XiaMen Univ., China) for insightful discussion, suggestions and technical supports in data post-processing. The DNS code provided by M. Gaudling is greatly appreciated.

* Lipo.Wang@sjtu.edu.cn

- [1] E. N. Lorenz, Deterministic non-periodic flow, J. Atmos. Sci. **20**, 130 (1963).
- [2] B. D. Hughes, Random walks and random environments (Oxford, 1995).
- [3] U. Frisch, Turbulence-the legacy of A.N. Kolmogorov (Cambridge, 1995).
- [4] B. Mandelbrot. The fractal geometry of nature (W.N. Freeman, New York, USA, 1983).
- [5] S. Corrsin, Random geometric problems suggested by turbulence. In Statistical Models and Turbulence. Lecture Notes in Physics, vol. 12 (ed. M. Rosenblatt C. Van Atta), (Springer, 1971) pp. 300-316.
- [6] A. E. Perry, and M.S.Chong, A description of eddying motions and flow patterns using critical-point concepts, Ann. Rev. Fluid Mech. **19**, 125 (1987).
- [7] L. P. Wang, and N. Peters. The length-scale distribution function of the distance between extremal points in passive scalar turbulence, J. Fluid Mech. **554**, 457 (2006).
- [8] L. P. Wang, and N. Peters. Length-scale distribution functions and conditional means for various fields in turbulence, J. Fluid Mech. **608**, 113, (2008).
- [9] L. P. Wang, and Y. X. Huang. Multi-level segment anal-

- ysis: definition and application in turbulent systems, J. Stat. Mech, 06018, (2015).
- [10] A. N. Kolmogorov. The Wiener spiral and some other interesting curves in the Hilbert space. Dokl. Akad. Nauk SSSR **26**, 115 (1940).
 - [11] B. Mandelbrot, and J. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Rev. **10**, 422 (1968).
 - [12] J. Beran. Statistics for Long-Memory Processes. Chapman Hall/CRC Monographs on Statistics & Applied Probability (Taylor & Francis, 1994).
 - [13] L. C. G. Rogers, Arbitrage with fractional Brownian motion, Mathematical Finance, **7**, 95 (1994).
 - [14] P. Doukhan, G. Oppenheim, and M. Taqqu. Theory and Applications of Long-Range Dependence (Birkhäuser Boston, 2002).
 - [15] A. Wood, and G. Chan, Simulation of Stationary Gaussian Processes in $[0, 1]^d$, J. Comput. Graph. Stat. **3**, 409 (1994).
 - [16] <https://www.ngdc.noaa.gov/mgg/global/>